

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-306076

(43)Date of publication of application : 02.11.2000

(51)Int.Cl.

G06T 1/00

G06F 12/00

H04N 1/00

(21)Application number : 11-112826

(71)Applicant : CANON INC

(22)Date of filing : 20.04.1999

(72)Inventor : HISAGAI MASAMI

(54) IMAGE PROCESSOR, CONTROL METHOD AND STORAGE MEDIUM

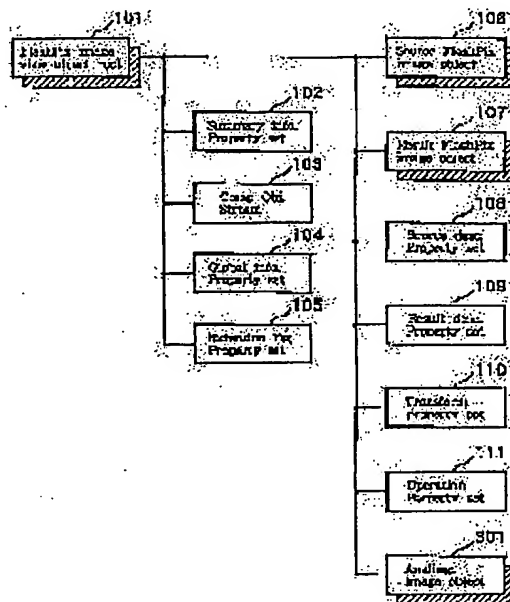
(57)Abstract:

PROBLEM TO BE SOLVED: To enable not only the retrieval of document images but also the provision of images with a desired definition.

SOLUTION: Character information and non-character information provided by inputting and analyzing an image including sentences are stored and managed.

At such a time, concerning an image part to be a non-character area, the image data of a hierarchical structure with different resolution are generated.

Then, the generated image data of the hierarchical structure are preserved in the unit of character information. Concerning the character information, an analyzed image object 301 is further provided in a flash pixel format and it is preserved here.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

含む複数の画像データを置くストレージを含む文書画像ファイルであって、前記画像データを文書解析した結果の文書解析データを含むことを特徴とする文書画像ファイルを記憶するメディア、または該文書画像データを通信経路を通して送受信する方法を提供する。

【0010】また、前記文書解析データには、前記画像データに関連する情報をリンクするためのアドレス情報を含むことを特徴とする。

【0011】

【発明の実施の形態】以下、添付図面に従って本発明に係る実施形態を詳細に説明する。

【0012】図1と図2は、FlashPix formatファイルの説明する図であり、先ず、これについて説明する。

【0013】ファイル内の各プロパティやデータには、MS-DOS（米国Microsoft社の商標）のディレクトリとファイルに相当する。ストレージとストリームによってアクセスする。図1、図2において、影付き部分がストレージを示し、影なし部分がストリームである。画像データや画像属性情報はストリーム部分に格納され、画像データは最小な解像度で階層化されており、それぞれその解像度の画像をSubImageと呼び、Resolution 0〜nで示してある。各解像度画像に対して、その画像を呼び出すために必要な情報がSubImage headerに、また画像データがSubImage dataに格納される。プロパティセットとは属性情報をその使用目的、内容に応じて分類して定義したもので、Summary Info、Property Set、Image Info、Property Set、Image Content Property Set、Extension list property Setがある。

【0014】[各プロパティの説明] Summary Info、Property Setは、FlashPix特有のものではなく、米国Microsoft社のストラクチャードストレージでは必須のプロパティセットで、そのファイルのタイトル・題名・著者・サムネイル画像等を格納する。

【0015】Image Content Property Setは、画像データの格納方法を記述する属性である。この属性には画像データの階層数、最大解像度の画像の幅、高さや、それぞれの解像度の画像についての幅、高さ、色の構成、あるいはJPEG圧縮を用いる際の量子化テーブル・ハフマンテーブルの定義を記述する。

【0016】Image Info、Property Setは、画像を使用する際に利用できる様々な情報、たとえば、画像がどのようなように取り込まれ、どのように利用可能であるかの情報を格納する。

【デジタルデータへの取込み方法/使いは生成方法に関する情報 (File Source)】

・画像の内容 (画像中の人物、場所など) に関する情報 (Content description)

・撮影に使われたカメラに関する情報 (Camera information)

・撮影時のカメラのセッティング (露出、シャッタース

50 【0023】さらに各解像度のレイヤの画像は64×6

ピード、焦点距離、フラッシュ使用の有無など) の情報 (Per Picture camera settings)

・デジタルカメラ特有解像度やモザイクフィルタに関する情報 (Digital camera characterization)

・フィルムカメラ名、製品名、種類 (ネガ/ポジ、カラー/白黒) などの情報 (Film description)

・オリジナルが写物や印刷物である場合の題名やサイズに関する情報 (Original document scan description)

・スキャン画像の場合、使用したスキャナやソフト、操作した人に関する情報 (Scan device)

Extension list property Setは上記FlashPixの基本仕様に含まれない情報を追加する際に使用する領域である。

【0017】図1の101〜111において、Flash Image View Objectは、画像を表示する際に用いるビューイングパラメータと画像データを合わせて格納する画像ファイルである。ビューイングパラメータとは画像の回転、拡大/縮小、移動、色変換、フィルタリングの処理を画像表示の際に適応するために記憶しておく処理情報のセットである。

【0018】Source/Result FlashPix Image ObjectはFlashPix画像データの本体であり、Source FlashPix Image Objectは必須、Result FlashPix Image Objectはオプションである。

【0019】Source FlashPix Image ObjectはOriginalの画像データを、Result FlashPix Image Objectはビューイングパラメータを使い画像処理した結果の画像を格納する。

【0020】Source/Result desc. Property setは上記、画像データの識別のためのプロパティセットであり、画像ID、変更禁止のプロパティセット、最終変更日時を格納する。

【0021】Transform property setは回転、拡大/縮小、移動のためのAffine変換係数、色変換マトリクス、コントラスト調整値、フィルタリング係数を格納している。次に画像データの取り扱いについて説明する。

【0022】図4に解像度の異なる複数の画像から構成される画像ファイルの例を示す。図4で最大解像度の画像がC×Rの画素で構成されており、その次に大きい画像はR/2×C/2であり、それ以降は横、縦、列とも1/2ずつ縮小し、列・行ともに64画素以下あるいは等しくなるまで繰り返す。このように階層化した結果、画像の属性情報として「1つの画像ファイル中の階層数」やそれぞれの階層の画像に対して、ヘッダ情報と属性データが必要となる。1つの画像ファイル中の階層の数や最大解像度の画像の幅、高さ、あるいはそれぞれの解像度の画像の幅、高さ、色構成、圧縮方式等に関する情報は前記Image Contents Property Setに記述される。

4画素のタイルに分割されている。画像の左上から順次64×64のタイルに分割すると、画像によっては右端および下端の一部に空白が生ずる場合がある。この場合はそれぞれ画素端および画素端画像を繰り返して挿入すること、64×64画像を格納する。FlashPix画像では、それぞれのタイル中の画像をJPEG圧縮、シグナルカラー、非圧縮のいずれかの方法で格納する。JPEG圧縮は、ISO/IEC、JTC1/SC29により国際標準化された画像圧縮方式であり、方式自体の説明はここでは割愛する。このようにタイル分割された画像データは、SubImage dataストリーム中に格納され、タイルの個数、個々のタイルのサイズ、データの開始位置、圧縮方法はすべてSubImage headerに格納されている。シングルカラーとは、前記1つのタイルがすべて同じ色で構成されている場合にのみ、個々の画素の値を記述することなく、そのタイルの色を1色で表現する方式である。この方法は特にコンピュータグラフィックスにより生成された画像に対して有効である。

【0024】図3は実施形態の特徴を示すものであり、FlashPix image view object rootに、Analyzed image objectを追加したものである。Analyzed image objectは、文書解析データであり、図5のようにBlock Information (501)、Text Data (502)、Key Data (503) から構成 *

```
struct header {
    int BInfo_offset; //Block Information dataのオフセット
    int BInfo_dataLang; //Block Information dataのデータ長 (単位: バイト)
    int BData_offset; //Block Dataのオフセット
    int BData_dataLang; //Block Dataのデータ長 (単位: バイト)
    int Key0_offset; //Key Dataのオフセット
    int Key0_dataLang; //Key Dataのデータ長 (単位: バイト)
} header;
```

という構造化形式である。

【0027】図7で、Block numberは、文書に含まれる領域の個数であり、図6の例では「5」(ブロック601〜606)である。Block DATAには、次の構造化 IDATAが先頭に配置されている。

```
struct tagBdata {
    int x1; //ブロックの左上角のx座標
    int y1; //ブロックの左上角のy座標
    int x2; //ブロックの右下角のx座標
    int y2; //ブロックの右下角のy座標
    int kind; //ブロックの種類
    int offset; //ブロックデータのオフセット
} BDATA;
```

ここで、ブロックの種類kindは、たとえば次のように定義する。

【0028】

0: テキスト

1: 表

2: 線画

*される。一般に、文書画像は、イメージスキヤナなどによって、取り込まれる。文書画像には、文字からなるテキスト領域 (テキストブロック)、表領域 (表ブロック)、線画領域 (線画ブロック)、画像領域 (ビクチャブロック) が含まれる。テキストブロックは、見出しテキストブロックと本文テキストブロックに識別することができる。文書画像から、これらの領域 (ブロック) を抽出し、そのブロックの位置・大きさ・種別を求めることができる。かかる技術としては特開06-068301号公報等がある。

【0025】たとえば、図6のように、文書画像601から、602、603、604、605、606の各ブロックが抽出され、該ブロックの位置・大きさ・種別が求められる。ブロック602は見出しテキストブロック、ブロック603および604は本文テキストブロック、ブロック605、606はビクチャブロックである。

【0026】また、図7は、Analyzed image objectの内容、すなわち文書解析データ (Block Information (501)、Text Data (502)、Key Data (503)) の表現方法の一例を示したものである。701は、Headerであり、その内容をC言語で表すと、

```
struct header {
    int BInfo_offset; //Block Information dataのオフセット
    int BInfo_dataLang; //Block Information dataのデータ長 (単位: バイト)
    int BData_offset; //Block Dataのオフセット
    int BData_dataLang; //Block Dataのデータ長 (単位: バイト)
    int Key0_offset; //Key Dataのオフセット
    int Key0_dataLang; //Key Dataのデータ長 (単位: バイト)
} header;
```

3: 画像

9: unknown

offsetで示される位置に、そのブロックのデータが記憶されている。たとえばテキストであれば、OCRされた結果の文字コードが記憶されている。また、画像であれば、画像のビットマップデータ、段であれば、たとえばRTFの表データ形式、線画であれば、IndyのGDIコマンド、またunknownならば画像のビットマップ形式が記憶されている。これらの記憶形式は、一例であって他の形式でも構わない。また、本実施形態では、このようなブロックのデータは解像度BDATAの直後に格納されているものとする。

【0029】ところで、この文書解析データ (Analyzed image object) は、FlashPix Imageobjectの通常の解像度の画像を、特開平06-068301号公報に開示された方法で作成される。あるいは、FlashPix image objectには無い他の解像度でスキャンされた画像から作成されても構わない。FlashPix image objectを作成するための適当なアプリケーションソフトウェアが必要な

50

ことはいずれまでもないが、本実施形態には直接関係ないので特に説明はない。このアプリケーションによって各プロパティ情報が入力されることが可能である。それとともに、Analyzed image object (Key Data) も入力することができ、Key Dataは、例えば、検索用のキーワード (該画像の特徴を現す言葉など) である。あるいは、Key Dataは、また本実施形態のもうひとつの特徴である。該画像データに関連する情報をリンクするためのアドレス情報、例えばインターネットのURL (Universal relocatable location) である。

[0030] 図8は、図3の画像を表示したアプリケーションの一例である。801はアプリケーションの親ウィンドウ、802は画像表示ウィンドウ、803はテキスト領域のテキストを表示するウィンドウ、804はKey Dataに記憶されているURLのリスト表示である。

[0031] 不図示のコンボボックス形式等のメニューを選択することにより、802の表示は異なる解像度の画像をFlashPix image objectから取り出して表示を変更できる。通常、表示装置はせいぜい100dpiの解像度であるので、表示用の画像としても100~200dpiが最適で、OCRのために使われてきた300~400dpiの画像を表示する必要はない。また、803では、300~400dpiの画像をOCRした結果のテキストが表示されている。不図示のスクロールバーを操作して、テキスト全体を見ることができ、また、804のURLのどれかをダブルクリックすると、インターネットエクスプローラやネットスケープなどのWebブラウザが起動して、そのページがロードされる。そのURLのページには例えば本画像についてのより詳しい説明が載っている。

[0032] 上記処理及び機能を実現するため、本実施形態においてファイリング装置に適用した場合の構成は図9の図になる。

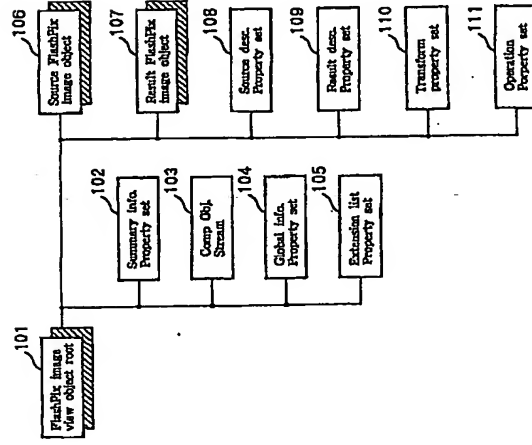
[0033] 図中、1は装置全体の制御を司るCPU、2はBIOSやブートプログラムを記憶しているROM、3は主記憶装置としてのRAMである。4はキーボードやマウス等の入力装置であり、5は原稿画像をカラー画像として読み取るスキャナである。6はネットワーク (例えばインターネット) に接続するネットワークインタフェースであり、7はビデオコントローラ (ビデオメモリを内蔵している)、8はビデオコントローラよりビデオ信号に基づいて像を表示する表示装置である。そして、9はスキャナ5で読み取った画像がデータバスとして格納する外部記憶装置 (たとえばハードディスク等) である。なお、外部記憶装置9には、OS及びアプリケーションシステムとして接続するアプリケーションも記憶されていて、電源投入時には、ROM2のブートプログラムにしたがって外部記憶装置9に格納されているOSが起動し、その後ファイリングアプリケーションを起動することになる。

ラムコードを読み出し実行することによって、達成されることは言うまでもない。この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記した記憶媒体は本発明を構成することになる。また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているオペレーティングシステム (OS) などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

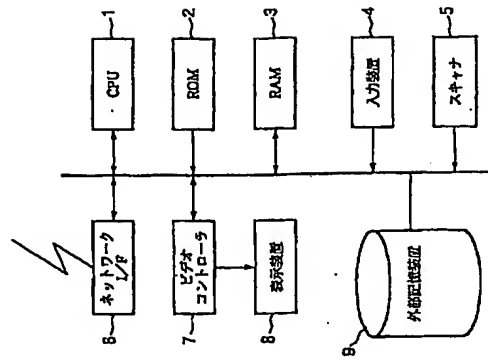
[0042] さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能記憶カードやコンピュータに接続された機能記憶ユニットに載るメモリに格納された後、そのプログラムコードの指示に基づき、その機能記憶カードや機能記憶ユニットに載るメモリなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

[0043] 以上説明したように本実施形態によれば、カラー画像を表示する解像度で表示でき、その画像のOCR結果も同時に表示ができる。また、Key Dataの

【図11】



【図9】



活用によりURLを記憶することにより、インターネットの関連ページへのリンクが可能となる。

[0044] 発明の効果 以上説明したように本発明によれば、文書画像の検索は勿論、要望される品位の画像を提供することが可能になる。

【図面の簡単な説明】

【図1】 FlashPix image object rootを示す図である。

【図2】 FlashPix image objectを示す図である。

【図3】 実施形態におけるFlashPix image object rootを示す図である。

【図4】 マルチ解像度での記憶を説明する図である。

【図5】 Analyzed image objectの図である。

【図6】 サンプリング画像を示す図である。

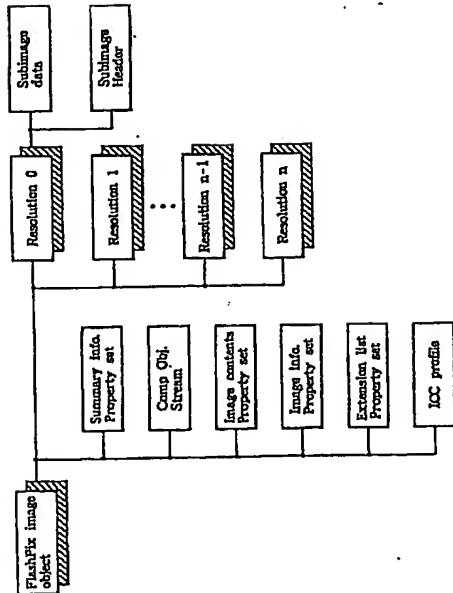
【図7】 Analyzed image objectの内容の表現方法の一例を示す図である。

【図8】 実施形態の画像フォーマットのアプリケーションの例を示す図である。

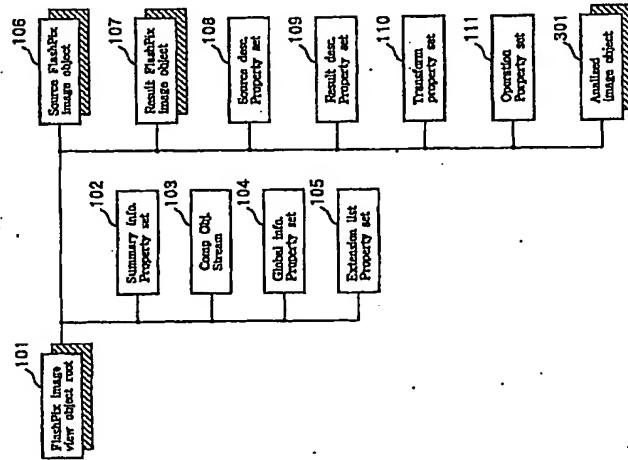
【図9】 実施形態が適用するファイリング装置のブロック構成図である。

【図10】 実施形態におけるファイリング処理手順を示すフローチャートである。

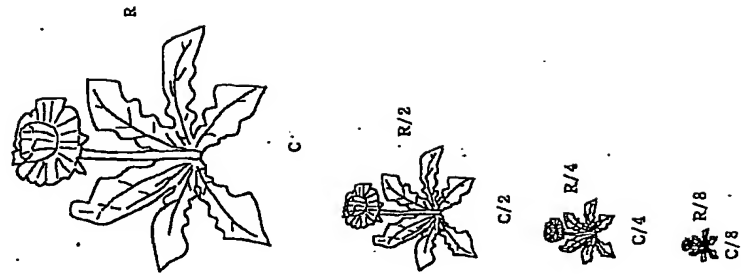
[圖 2]



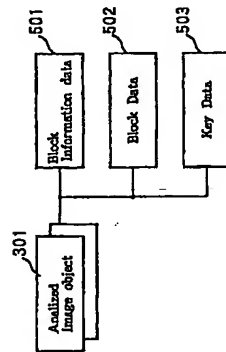
[圖 3]



[圖 4]



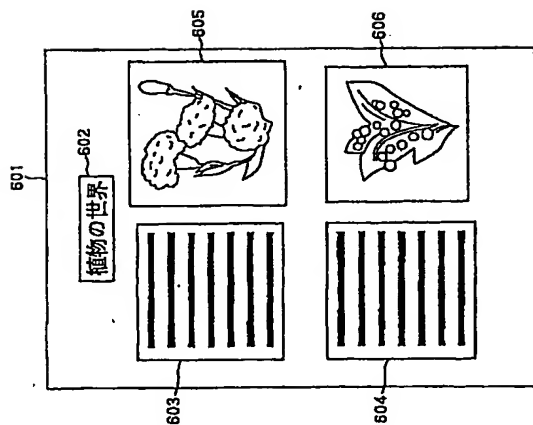
[圖 5]



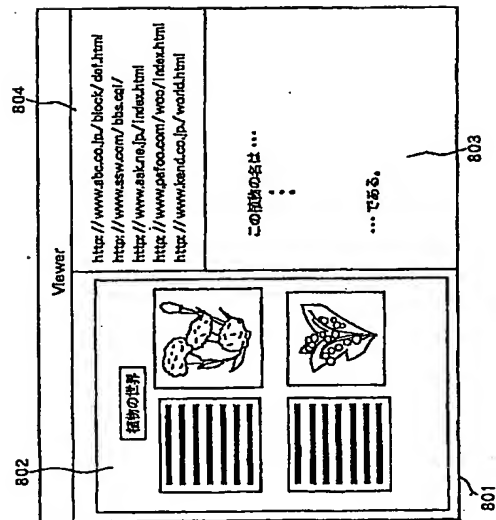
[圖 7]

Header	701
Block number	702
Block information data (1)	703
...	...
Block information data (n)	704
Block information data (a)	705
Block Data (1)	706
...	...
Block Data (a)	707
Key Data	708

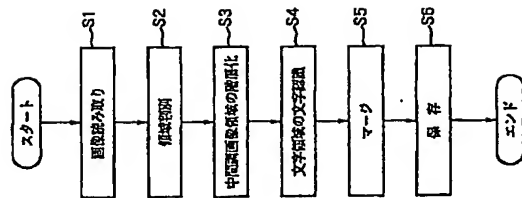
【図6】



【図8】



【図10】



フロントページの続き

Fターム(参考) 58050 BA06 BA10 BA16 CA04 CA07
DA06 EA03 EA05 EA06 EA10
EA12 EA19 FA02 FA12 FA19
CA08
58082 AA13 EA01
5C062 AA06 AB17 AB23 AB38 AB42
AC08 AC22 AC24 AC51 AF00
BA00